Limnol. Oceanogr., 45(2), 2000, 408–418

The complete genomic sequence of the marine phage Roseophage SIO1 shares homology with nonmarine phages

Forest Rohwer[1]
Marine Biology Research Division, Scripps Institution of Oceanography, San Diego, California 92037

Anca Segall
Department of Biology and the Microchemical Core Facility, San Diego State University, San Diego, California 92182

Grieg Steward
Monterey Bay Aquarium Research Institute, Monterey, California 95039

Victor Seguritan
Department of Biology and the Microchemical Core Facility, San Diego State University, San Diego, California 92182

Mya Breitbart
Marine Biology Research Division, Scripps Institution of Oceanography, San Diego, California 92037

Felise Wolven
Department of Biology and the Microchemical Core Facility, San Diego State University, San Diego, California 92182

Farooq Azam
Marine Biology Research Division, Scripps Institution of Oceanography, San Diego, California 92037

Abstract

Viruses are ubiquitous components of the marine environment, frequently reaching concentrations of 10^7–10^8 viruses per milliliter of surface seawater. The majority of these viral particles are bacteriophages (phages). Although the oceans are probably the largest pool of bacteriophages on the planet, the evolutionary relationships of marine phages to phages from other environments are unknown. To address this issue, we have completely sequenced the genome of the lytic marine phage, Roseophage SIO1, that infects the heterotrophic marine bacterium *Roseobacter SIO67*. This phage has an isometric capsid with a diameter of approximately 43 nm, a short tail, a buoyant density of 1.49 g cm^{-3} in CsCl, and a 39,906-bp dsDNA genome. Sequence similarities and relative positions within the genome suggest that three of the open reading frames (ORFs) are homologous to the primase, DNA polymerase, and endodeoxyribonuclease I proteins of coliphages T3 and T7. The results are consistent with the mosaic theory of phage evolution and indicate a genetic link between marine and nonmarine phages. Additionally, basic life histories of marine phages can be elucidated by comparison of complete genomes to those of other extensively studied phages (e.g., lambda, T4, T7). The DNA replication machinery of Roseophage SIO1 shows a clear homology with that of coliphages T3 and T7, suggesting that the process of DNA replication may be similar among these phages. The Roseophage SIO1 genome also encodes four predicted proteins involved in phosphate metabolism (RP PhoH, RP ribonucleotide reductase, RP Thy1, and RP endodeoxyribonuclease I) suggesting that phosphate recycling is important to Roseophage SIO1's life cycle. Other interesting clues about Roseophage SIO1's life history come from the absence of certain expected protein regions. For example, we have not been able to identify the Roseophage SIO1 structural proteins (e.g., capsid proteins) by homology to other phages. It is also conspicuous that the Roseophage SIO1 genome lacks a recognizable RNA polymerase, an essential component of T3 and T7 life cycles. Analysis of the Roseophage SIO1 genome shows that marine and nonmarine phages are genetically related but basic life histories may be significantly different.

Approximately one half of oceanic primary production is assimilated by marine heterotrophic bacteria and the fates of these bacteria greatly influence global biogeochemical processes in the ocean (reviewed in Azam 1998). Grazing by protists is expected to transfer bacterial-derived energy and biomass into higher trophic levels (Sherr and Sherr 1984; Fenchel 1986). In contrast, bacterial lysis by phages results in production of dissolved organic matter that can be utilized by other heterotrophic bacteria (reviewed in Fuhrman 1999). Discovery of large numbers of viruses in marine waters and the high incidence of virally infected marine bacteria suggested that lysis might be a major fate of bacteria (Bergh et al. 1989; Proctor and Fuhrman 1990). It is currently estimated that in surface marine waters 10–20%, and sometimes

[1] Corresponding author (forest@ucsd.edu).

Acknowledgments

The authors would like to thank Roger Hendrix and Dennis Bamford for helpful hints about shotgun sequencing, phage evolution, and electron microscopy. Thank you to Jóse González and Mary Ann Moran (University of Georgia) for the different *Roseobacter* isolates. This work was supported by NSF grants OCE99-00301, OCE93-12695, OPP95-30851, and by the SDSU-CSUPERB Microchemical Core Facility.

over 50%, of heterotrophic bacterial production is converted to dissolved organic matter by phage lysis (Steward et al. 1992; Fuhrman and Suttle 1993; Suttle 1994; Fuhrman and Noble 1995; Steward et al. 1996).

Besides being important contributors to worldwide marine biogeochemical processes, marine phage are also assumed to be important in structuring the ocean's microenvironments. Phage are relatively host specific (Goyal et al. 1987; Moebus and Nattkemper 1991; Moebus 1992; Suttle and Chan 1993,1994; Waterbury and Valois 1993; Kellogg et al. 1995); therefore, individual bacterial species in a community will be affected differentially by the surrounding phage community. This process has been suggested as a mechanism that increases the overall microbial diversity (Levin et al. 1977; Fuhrman and Suttle 1993; Waterbury and Valois 1993; Miller 1998). Phage can affect the growth rate of specific bacterial species by selecting for resistant but slow growing strains, presumably by forcing the downregulation of viral receptors (e.g., siderophores, flagella, permeases) that are also necessary for rapid growth (Levin et al. 1977; Lenski 1988; Bohannan and Lenski 1997). Virus-mediated lysis can also affect the growth rates of noninfected populations via nutrient regeneration (Middelboe et al. 1996). Moreover, phage may influence marine microenvironments through the release of enzymes during bacterial lysis. The released enzymes may transform particulate organic matter to dissolved organic matter without associated respiration (Smith et al. 1992).

Marine phages morphologically resembling the Myoviridae, Siphoviridae, and Podoviridae families have been observed using electron microscopy (reviewed in Børsheim 1993). The majority of the native marine phages observed have capsid sizes of 30–60 nm and approximately half have recognizable tails (Børsheim 1993). Analysis of marine phage genomes by pulse field gel electrophoresis suggests that most of the marine viral genomes are between 20 and 70 kb in length (Wommack et al. 1999; Steward and Azam in press). Phages with RNA genomes appear to be rare in the marine environment (Hidaka and Ichida 1976; Børsheim 1993). Marine phage diversity is believed to be very high (Fuhrman 1999). Moebus and Nattkemper (1991) and Moebus (1992) screened over 900 culturable marine prokaryotes and found that approximately a third were susceptible to infection by at least one, and often several, lytic phage. In these studies, the methods underestimated the positives and the authors concluded that the majority of culturable bacterial strains were probably susceptible to phage infection.

Other studies have approached the question of phage diversity by investigating the number of different phage that infect one species of bacteria. Kellogg et al. (1995) isolated 60 phage from Florida and Hawaii that infected *Vibrio parahaemolyticus* and analyzed them by restriction fragment length polymorphism, host specificity, and Southern blotting. Restriction fragment length polymorphism analysis separated the 60 isolates into six distinct groups. The genetic relatedness of these six groups was shown by hybridization of a 1.5-kb DNA probe originally cloned from one of the isolates. None of the 60 isolates was able to infect closely related *Vibrio* species. Studies with the *Synechococcus* have shown that there are a large number of cyanophages that differ in their host specificity and morphology (Suttle and Chan 1993, 1994; Waterbury and Valois 1993). Together, these results strongly suggest that phage diversity is at least as high, and probably higher, than the diversity of marine microbes (i.e., there is at least one, and usually more, phage that infects each marine microbe that has been studied and each phage is usually specific to a single microbial species or strain). Since the number of prokaryotic species in the ocean is estimated to comprise several million species, there should be millions of different marine phage (Hawksworth 1995; Furhman and Campbell 1998).

Although the abundance and diversity of marine phages have been studied, the evolutionary relationships of marine phages to other phages, and to each other, are essentially unknown. To begin addressing this question we have determined the complete genomic sequence of Roseophage SIO1, a lytic phage that infects the marine heterotrophic bacteria *Roseobacter SIO67*. Analysis of the Roseophage SIO1 genome revealed a portion of the genome that encodes DNA replication proteins (i.e., DNA polymerase, primase, and endodeoxyribonuclease I) that are evolutionarily related to those found in the coliphages T3 and T7.

Materials and methods

Isolation of Roseobacter *SIO67 and Roseophage SIO1*—One milliliter of seawater collected from the end of the pier at Scripps Institution of Oceanography on 30 October 1989 was spread on ZoBell 2216E agar plates (5 g peptone, 1 g yeast extract, 0.01 g $FeC_6H_5O_7$, and 1.5 g agar added per liter of 80%, 0.45 μm-filtered seawater) and incubated at 16°C for 3 d. Individual colonies were picked and restreaked twice. A very small, nondescript colony was designated SIO67, inoculated into 30 ml of ZoBell 2216E medium, and incubated at 16°C with shaking until turbid (several days). To prepare cells for plating, the culture was centrifuged for 10 min at 5,000 rpm in a clinical centrifuge and the bacterial pellet resuspended in 7.5 ml of 0.45 μm-filtered, autoclaved seawater. A bacterial lawn was prepared using the agar overlay technique by adding 3 ml of molten top agar (0.7% agar in ZoBell; 50°C) to 200 μl of cells and pouring over plates. A seawater sample to be screened for bacteriophage was collected on 3 November 1989, treated with 0.3% chloroform (vol : vol), and stored at 4°C for several weeks. A 10-μl aliquot of the chloroform-treated seawater was spotted on the host cell lawn and incubated at 16°C. Growth of a dense bacterial lawn required approximately a week, and a plaque appeared between 7 and 12 d. Bacteriophages were recovered by picking the plaque with a sterile Pasteur pipette and eluting in 0.5 ml of MSM (32.5 mM NaCl, 12 mM $MgSO_4$, 50 mM Tris, 0.1% gelatin) at 4°C overnight. Dilutions of eluted phages were combined with host cells and plated by agar overlay to obtain fresh plaques. The picking and replating of individual plaques was serially repeated 3× to insure isolation of a single clonal phage population. A phage stock was prepared by adjusting phage concentration to give confluent lysis on agar overlay plates (3,000–5,000 PFU per plate). MSM was added to the plates and phages were eluted with rocking at 4°C overnight. Eluates were pooled, filter sterilized (0.2 μm Acrodisc; Gelman), and stored at 4°C.

Electron microscopy of Roseophage SIO1—Aliquots from plate lysates were left unfixed or were fixed by addition of glutaraldehyde (electron microscopy grade; Ted Pella) to 1% final concentration. Drops of the bacteriophage suspensions in MSM were adsorbed to carbon-stabilized formvar supports on 200-mesh copper grids (Ted Pella) for several minutes. Excess liquid was then removed by touching the edge to adsorbent paper. Fixed bacteriophage were also deposited by ultracentrifugation (35,000 rpm for 1 h in a Beckman SW41 rotor) onto grids rendered hydrophilic by glow discharge (Suttle 1993). Grids were rinsed once by immersion in a drop of purified H_2O. Adsorbed phage were negatively stained by adding a drop of 2% phosphotungstic acid (for the unfixed sample) or 2% uranyl acetate (for the fixed sample). For the unfixed sample, excess stain was immediately wicked away and the grid allowed to air dry. The fixed sample was stained for 20–25 s prior to wicking away excess stain and air-drying. Samples were examined by electron microscopy at 80 kV accelerating voltage and photographed at ×200,000 magnification. A calibration standard consisting of latex beads on a waffle grating replica (Ted Pella) was also photographed to calibrate microscope magnification. Electron microscopic negatives were overhead-projected onto paper and the diameters of five separate viruses from four negatives were measured. Three edge-to-edge and three vertex-to-vertex measurements were also made for each virus.

Sequencing of Roseobacter *16S rDNA gene fragment*—Chromosomal DNA was prepared from 1.5 ml of a dense culture of bacterial isolate SIO67 (Sambrook et al. 1989). A ribosomal gene fragment was polymerase chain reaction (PCR) amplified (1× PCR buffer [PE Applied Biosystems, Inc.; Foster City, California], 10 μM nucleotides, 1 mM $MgCl_2$, 0.2 μM of each primer, and 300 ng DNA template; 30 cycles of 94°C for 1 min, 57°C for 1 min, 72°C for 1 min) using a biotinylated forward primer 50–68F (eubacterial) and a reverse primer 536–519R (universal; Schmidt et al. 1991). PCR products were purified and complementary strands separated using streptavidin-coated magnetic beads (DYNAL) according to the manufacturer's instructions. Each strand was sequenced using dye-terminator technology (PRISM Dyedeoxy Terminator kit) and analyzed on an automated sequencer (PE Applied Biosystems, Inc.) according to the manufacturer's instructions.

Analysis of Roseophage SIO1's ability to infect other Roseobacter—The 13 other marine *Roseobacter* species or strains (d22, dnt3, ddnt1, ddnt8, GAI-5, GAI-16, GAI-37, GAI-101, GAI-109, GAI-111, and EE-36) and closely related bacteria (*Saguttula stellata* E-37, and *Sulfitobacter pontiacus* ChLG 10) were generously provided by Jóse González and Mary Ann Moran (University of Georgia) (Gonzalez and Moran 1997). Exponentially growing cultures of these bacteria were incubated with a Roseophage SIO1 lysate ($\approx 10^{12}$ PFU μl^{-1}) for 30 min and then plated in ZoBell top agar. (Note: During the initial isolation of *Roseobacter SIO67* and Roseophage SIO1, the ZoBell 2216E formulation noted above was used; for all subsequent culturing, the ZoBell medium contained 0.5% peptone, 0.1% yeast extract, filtered seawater, 1.5% agar.) The plates were monitored for plaque formation. Different concentrations (i.e., ½ dilutions) of each strain were used to ensure that the initial concentration of bacteria did not bias the results.

Isolation of Roseophage SIO1 DNA—A single colony of *Roseobacter SIO67* was inoculated into 50 ml ZoBell media and incubated at 18°C for approximately 40 h with slow shaking. This culture was then transferred to 500 ml fresh ZoBell media and incubated for an additional 4 h. The cells were pelleted and resuspended into 50 ml of MSM, and the cell concentration was determined by spectrophotometry (2×10^9 *Roseobacter SIO67* cells $ml^{-1} \times 1$ $O.D._{600}$). A total of 1×10^{11} *Roseobacter SIO67* cells were harvested and incubated with 1,000 Roseophage SIO1 for 30 min at room temperature. The phage–host mixture was then poured into 1 liter of ZoBell media and incubated at room temperature until significant bacterial lysis was observed (approximately 13 h).

Phage were purified essentially as described by Sambrook et al. (1989). Ten milliliters of chloroform was added to each liter of phage lysate and incubated at room temperature for 15 min. ribonuclease (RNase) A (10 μg ml^{-1}) and deoxyribonuclease (DNase) 1 (0.02 SU ml^{-1}) were added to the liquid lysate and incubated at room temperature for 1 h. Sodium chloride was added to a final concentration of 1 M, and the mixture was incubated on ice for 30 min. Cellular debris was removed by centrifugation at 9,000 rpm in a Sorvall GSA rotor for 10 min, and then the supernatants were filtered through one layer of KimWipe tissue paper. Polyethylene glycol (PEG 8000; 100 g L^{-1}) was added, and the mixture was incubated overnight at 4°C. The PEG–phage precipitates were collected by centrifugation at 9,000 rpm using a Sorvall GSA rotor for 30 min. The pellets were resuspended into 5 ml MSM, and the PEG was extracted using chloroform. Cesium chloride (0.75 g ml^{-1}) was added to the concentrated phage and transferred to ultracentrifuge tubes (Beckman, Fullerton, California). Centrifugation was carried out in a SW41 rotor at 35,000 rpm ($110,000 \times g$) for 24 h at 4°C, and the purified phage band was removed. The buoyant density of the phage was estimated by weighting a known volume of the harvested band using a micropipettor and an analytical balance (buoyant density = weight in grams per volume in cm^3). Phage DNA was extracted by adding a 0.1 volume of 2 M Tris·HCl (pH 8.5)/0.2 M EDTA and 1 volume formamide. After incubating for 30 min at room temperature, two volumes of ethanol (100%) were added to this mixture to precipitate the DNA. The DNA was collected by centrifugation and washed with 70% ethanol twice. Finally, the DNA pellet was resuspended into water and the concentration was determined by spectrophotometry. This procedure yielded approximately 50 μg of Roseophage SIO1 DNA from each liter of lysate.

Construction of the shotgun library—Random fragments of Roseophage SIO1 DNA were produced by DNase I digestion in the presence of manganese chloride. Eight μg of DNA was resuspended into 50 mM Tris·HCl (pH 7.6), 10 mM $MnCl_2$, 100 μg ml^{-1} bovine serum albumin, and 0.01 SU ml^{-1} DNase I. Aliquots were removed 0, 0.5, 1, 2, 5,

10, 15, 30, and 60 min after addition of the digestion mixture and immediately transferred to a tube containing Tris-buffered (pH 7.0) phenol. After a phenol : chloroform (1 : 1) and two chloroform extractions, the fragmented DNA was precipitated, pelleted, washed with 70% ethanol, and dried. The fragmented DNA was resuspended in 23 μl of Blunt-ending Mix (100 μM dNTPs, 1× T4 DNA Pol Buffer (New England Biolabs [NEB]; Beverly, Massachusetts) and heated at 65°C for 30 min to resuspend DNA and inactivate any DNase I that was carried over. After cooling to room temperature, 2.5 U Klenow fragment (NEB) and 5 U T4 DNA polymerase (NEB) were added, and the reaction was incubated at 37°C for 1 h.

The fragmented and blunt-ended phage DNA was separated on a 1% agarose/TAE (tris acetate EDTA buffer (40 mM tris acetate, 2 mM EDTA)) gel and fragments in the 500–1,000-bp range were excised. The DNA fragments were isolated from the gel using the UltraClean Gel Spin kit (MolBio; Solana Beach, California) and ligated into pCR–Blunt (Invitrogen; San Diego, California) according to the manufacturer's protocol. The ligation was heat-shock transformed into TOP10 cells (Invitrogen) and spread onto LB/Kanamycin plates.

Preparation of PCR fragments for sequencing—Individual colonies were picked from the plates and streaked into the bottom of 0.5 ml PCR tubes. Fifty microliters of the PCR master mix (10 mM Tris·HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl, 200 μM dNTPs, 0.2 U Taq DNA polymerase [Boehringer Mannheim]) were added to each tube and overlaid with 90 μl of light mineral oil. Amplification was carried out as follows: 94°C for 5 min (1 cycle); 30 cycles of 94°C for 1 min, step down annealing temperature starting at 65°C and decreasing 1°C every two cycles, an extension step of 72°C for 3 min; and a final 5-min extension step at 72°C. Primers were removed using UltraClean PCR Clean-Up kit (MolBio).

Sequencing and assembly of fragments—Automated DNA sequencing was carried out using ABI PRISM BigDye Terminators (PE Applied Biosystems, Inc.) on an ABI377XL sequencer. Inserts from the shotgun libraries were sequenced using the M13 forward or reverse primers. Sequences were assembled into continuous sequences (contigs) using the Gene Codes Sequencher 3.0.6 software. After fourfold coverage of the genome, the remaining gaps were closed in one of two ways. Initially, a PCR reaction with a mixture of primers designed to the ends of the contigs was performed using purified phage DNA as a template. The resulting PCR products were separated on an agarose gel, excised, purified, and sequenced. Although this method closed many gaps, ultimately it was not a very efficient approach. Eventually, we resorted to ''primer-walking'' using purified phage DNA.

Computer software and analysis—Sequencher (Gene Code Corp.; Ann Arbor, Michigan) was used to assemble fragments of the genome into contigs. Open reading frames (ORFs) were identified visually by scanning the genome for relatively large regions without stop codons. Possible homologs to Roseophage SIO1 proteins were identified by checking predicted protein translations of each ORF against GeneBank using Psi(Ψ)-BLAST (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-psi_blast) (Altschul et al. 1997). Phylogenetic trees for the 16S rDNA, primase, DNA polymerase, and endodeoxyribonuclease I were constructed by aligning sequences using Multalin version 5.3.3 (Corpet 1988), trimming sequences so that only the most conserved regions remained (this was done using a Perl script available from V. S. upon request), and analyzing alignments using various algorithms found in PHYLIP (Phylogeny Inference Package; version 3.5c; distributed by J. Felsenstein, at http://evolution.genetics.washington.edu/phylip.html).

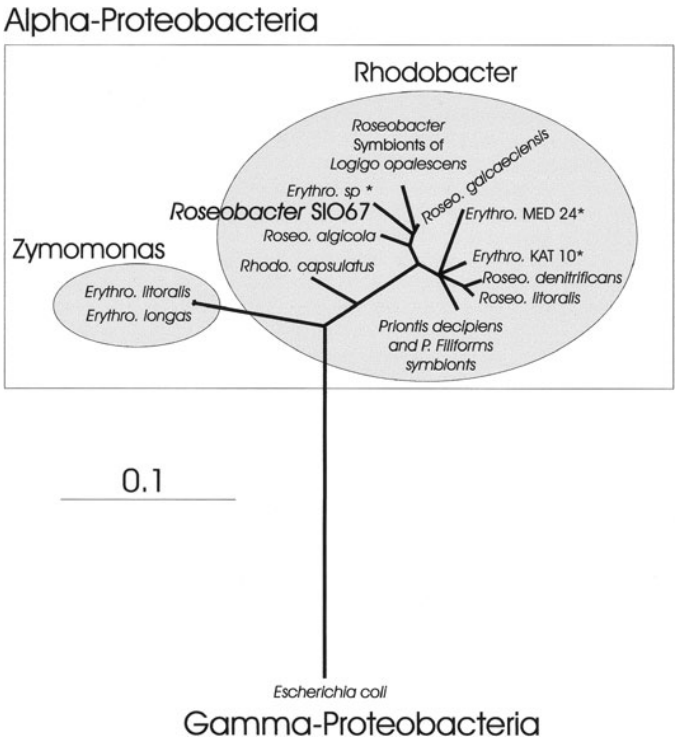

Fig. 1. Phylogeny of *Roseobacter SIO67*. Asterisk indicates isolates designated as *Erythrobacteria* by Pinhassi et al. (1997).

PCR to detect Roseophage SIO1 in seawater and MSM samples—PCR primers to an intergenic region of Roseophage SIO1 were designed to positions in ORFs 4 (5′-GGTTGCCGTTAAGCCTAG-3′) and /5 (5′-GTTTCAGTCTGACTGCATCG-3′). Fifty reactions, with 1 μl of the sample in MSM or seawater, were assembled and amplified using the same conditions as described above. The sensitivity of the PCR was tested by using various dilutions of Roseophage SIO1 in MSM to develop a standard curve. To check the various Roseophage isolates, plaques were placed into 50 μl MSM, and vortexed, and then 1 μl was titered. In all cases, the 1-μl sample contained >1,000 phage. Unless otherwise noted, all chemicals were purchased from Sigma Chemical Company (St. Louis, Missouri).

Results

The phage–host system used in this study was isolated in 1989 on ZoBell 2216E, a common medium for marine het-

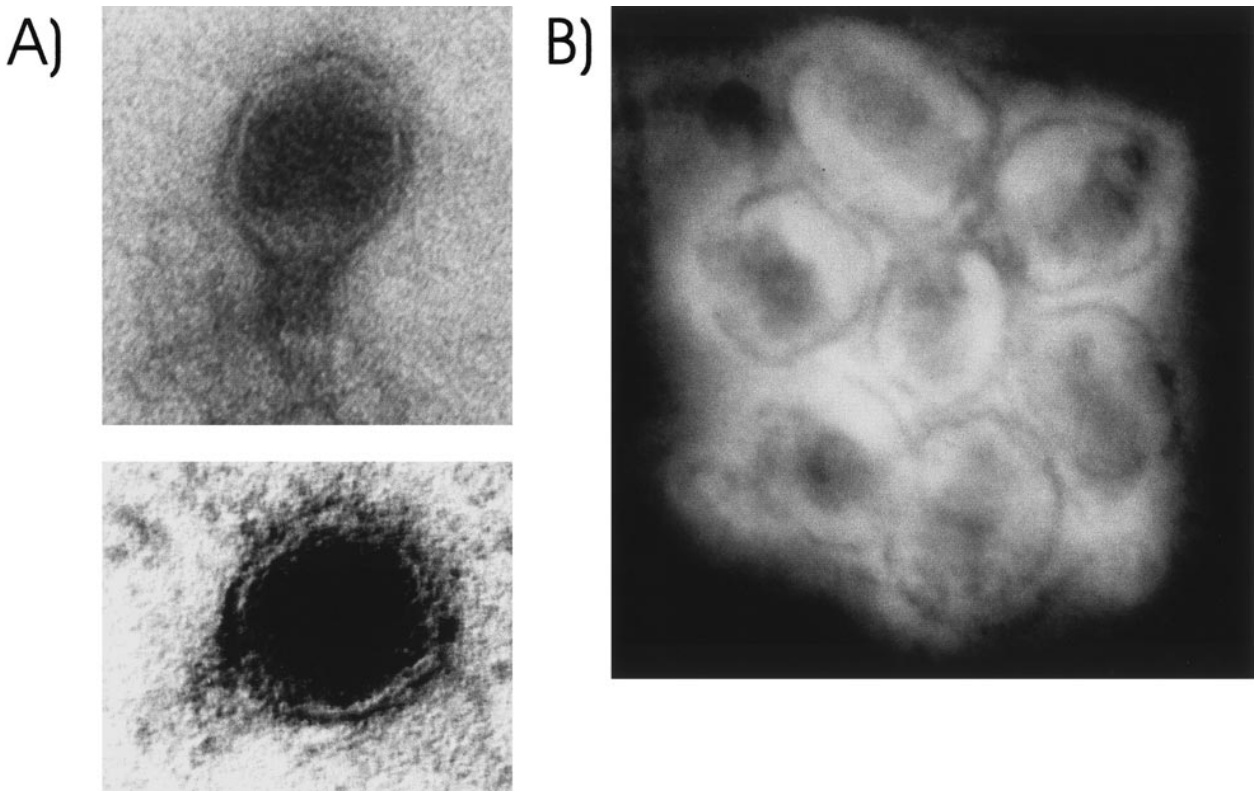

Fig. 2. Electron micrographs of Roseophage SIO1. (A) Fixed in 1% glutaraldehyde and stained with 2% uranyl acetate and (B) unfixed and stained with 2% phosphotungstic acid. ×100,000.

erotrophic bacteria. Analysis of the host's 16S rDNA sequence (GenBank accession number AF190747) showed that it is identical to a marine bacterium isolated by Pinhassi et al. (1997). Although these authors classified their isolate as an *Erythrobacter* sp., subsequent additions to the rDNA database suggest that this bacterium is more closely related to the *Roseobacter* genus (Fig. 1). Consequently, the host was named *Roseobacter SIO67*. Roseophage SIO1 is a lytic phage that infects *Roseobacter SIO67*.

Roseobacter sp. are common marine isolates and can constitute a significant proportion of the total heterotrophic bacterial community (Gonzalez and Moran 1997). To test the host specificity of Roseophage SIO1, 13 closely related *Roseobacter* isolates were incubated with the phage and monitored for plaque formation in top agar, indicating a productive infection. Roseophage SIO1 was found to exclusively lyse the original host *Roseobacter SIO67*.

Electron micrographs of Roseophage SIO1 fixed with 1% glutaraldehyde and stained with 2% uranyl acetate revealed an isometric capsid with a hexagonal cross section (Fig. 2A). A short tail is present. The phage capsid diameter is 43 ± 0.8 nm from edge to edge and 47 ± 1 nm from vertex to vertex (mean ± SD, $n = 15$). Tails were often poorly delineated and difficult to measure accurately. However, a single measurement from a photograph indicates a tail length of roughly 20 nm. Negatively stained viruses prepared without fixative show the effects of osmotic shock; the capsid loses its hexagonal shape, becoming amorphic, and a bullet-shaped core is visible (Fig. 2B). This core is reminiscent of the nucleocapsid cores of coliphages T3 and T7 (Calendar 1988). Electrophoresis (both conventional and pulse field gel) of the phage genome revealed a band 36–37 kb in size (data not shown; Wikner 1993). The Roseophage SIO1 genome was also analyzed on alkaline gels to check for single-stranded DNA. No evidence of single-stranded DNA was found (data not shown). The phage genome was completely degraded by DNase I treatment, but was insensitive to RNases A or T1 (data not shown). Additionally, the genome could be cut by restriction enzymes such as *Alu*I and *Sau*3AI that are only active on double-stranded DNA (dsDNA; data not shown). Therefore, it was concluded that Roseophage SIO1 has a dsDNA genome. Pretreatment of the phage with chloroform did not significantly reduce infectivity, indicating that a lipid envelope is not required for phage viability. Finally, the phage has an approximate buoyant density of 1.49 g cm^{-3}. Based on these physical characteristics, Roseophage SIO1 was identified as a member of the Podoviridae family of phage (Murphy et al. 1995).

The complete genome of Roseophage SIO1 was determined by a combination of shotgun cloning and direct sequencing. The sequence has been submitted to GenBank and assigned the accession number Roseophage SIO1 AF189021. The genome is 39,906 bp long and has a GC content of 46.2%. Potential open reading frames (ORFs) were identified by scanning the sequence for regions greater than 240 bp long with fewer than two stop codons. These criteria would identify ORFs that encoded small proteins (i.e., 80 aa), as well as identifying ORFs that are translated

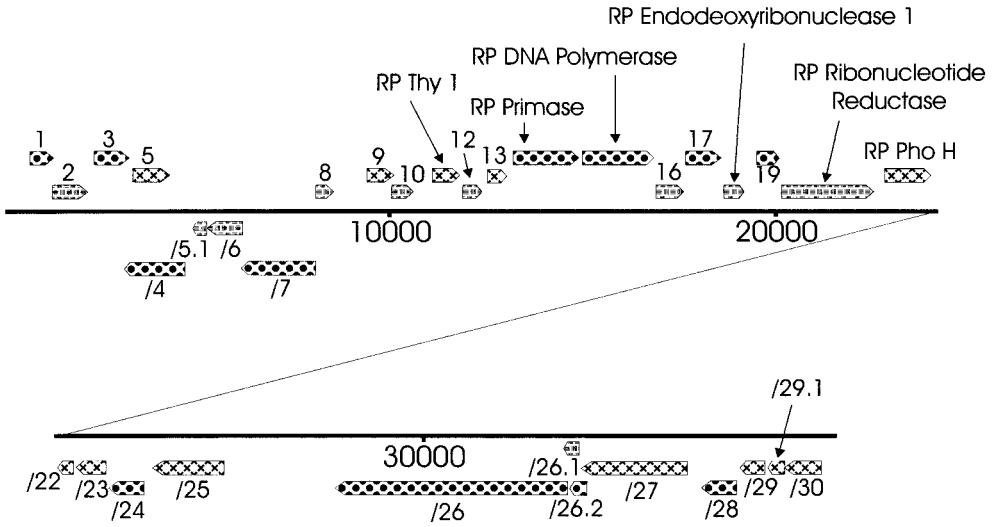

Fig. 3. Genomic organization of Roseophage SIO1. The genome was oriented such that the plus strand (+) of the primase, DNA polymerase (DNA pol), and endodeoxyribonuclease 1 ORFs read 5′ to 3′ starting on the left. The 6 ORFs encoding proteins homologous to previously predicted proteins were named after the homolog with the prefix RP (Roseophage). Numbers were assigned sequentially from the left for those ORFs without significant homology to other proteins. ORFs that are translated in the opposite direction have a slash (/) before the designating number.

via ribosomal slippage. For this analysis, it was assumed that Roseophage SIO1 utilizes the universal translational code for bacteria. Potential ORFs were further scrutinized for the presence of a 5′ start codons (e.g., ATG) and a Shine-Dalgarno sequence. Using this approach, a total of 30 ORFs were identified. The ORFs were translated in silico and submitted to the National Center for Biotechnology Information's (NCBI) PSI-BLAST program for identification of homologs. The genome also contains two identical inverted repeats at positions 50–250 and 36,668–36,868. These repeats could be involved in replication of Roseophage SIO1.

PSI-BLAST analysis revealed a region of the genome that encodes proteins similar to primase, DNA polymerase (DNA pol), and endodeoxyribonuclease I of coliphages T3 and T7. The Roseophage SIO1 genome was oriented such that the plus strand (+) of these three ORFs reads 5′ to 3′ from left to right. All subsequent numbering of the genome followed this orientation (Fig. 3). A total of six ORFs encoding proteins homologous to previously predicted proteins were identified (Table 1). In these cases, the ORFs were named after the homolog with the prefix RP (e.g., the homolog to the DNA polymerase of coliphage T3 was named RP DNA Pol). Those ORFs without significant homology to other proteins were sequentially numbered starting from the left side of the genome. ORFs that are translated in the opposite direction are designated by slash (/) before the designating number. Table 1 lists the name and coordinates of each ORF, as well as the name and accession number of the most similar homolog.

To determine the evolutionary relationships of Roseophage SIO1–predicted proteins to other homologous proteins, phylogenetic analyses were performed. Since the overall evolutionary relatedness of phage to each other and other organisms is unknown, cladograms incorporating homologs of the Roseophage proteins from Bacteria, Archaea, and Eukaryotae were produced. Because this approach is prone to producing long branch attraction artefacts, multiple phylogenetic reconstruction algorithms were tested, including: (1) the distance matrix algorithm FITCH (an additive tree model that does not assume an evolutionary clock), (2) KITSCH (an ultrametric model that assumes an evolutionary clock), (3) NEIGHBOR (a distance matrix method that produces an unrooted tree without the assumption of a molecular clock), and (4) a maximum parsimony model. In addition, consensus trees were produced. Of these approaches, KITSCH produced trees that most accurately represented the appropriate evolutionary relationships of the bacterial, archaeal, and eukaryotic proteins. Therefore, the trees shown in Fig. 4 were results of KITSCH analyses. For each of the three proteins, Roseophage SIO1 clusters with the coliphages T3 and T7. Although the cladograms do not group all of the phage proteins together in a clade, it is interesting to note that these phage proteins do not group closely with those of their hosts.

The other recognizable Roseophage SIO1 encoded proteins RP Thy1, RP reductase, and RP PhoH are not related to coliphage T3- or T7-encoded proteins. Although ribonucleotide reductases have been reported in a number of other phages (e.g., mycobacteriophages L5 and D29; Table 1; Calendar 1988; Hendrix et al. 1999), no homologs to Thy1 and PhoH have been reported in any other phage.

In order to examine Roseophage SIO1 in the environment, samples from the Puget Sound and Scripps Pier were screened for other phage that lytically infect *Roseobacter SIO67*. The Scripps Pier samples were taken 8 yr after the original isolation of Roseophage SIO1. Several plaques containing lytic phage that infect *Roseobacter SIO67* were purified from these samples. PCR primers were then designed to amplify an intergenic, noncoding region of Roseophage

Table 1. Roseophage SIO1 ORFs. The highest ranking match revealed by the initial round of PSI-BLAST analysis is given for the appropriate ORFs. The accession number of the most significant match precedes the description of the protein. The percentage of amino acid identities and positives (between the predicted Roseophage SIO1 sequence and the match) are also listed. Significant ($P < 0.005$) and nonsignificant hits to other phage are also listed. See text for further explanation.

ORF	Coordinates	Matches	Identity (%)	Positive (%)
1	441–851			
2	778–1521	Possible hit to minor structural protein of bacteriophage sfi11		
3	1581–2429			
/4	2446–4221			
5	2531–3346			
/5.1	4695–5261			
/6	5670–6419	Possible hit to phage ctx protein of bacteriophage phi ctx of *Pseudomonas aeruginosa*		
/7	6355–7977			
8	7891–8202			
9	9494–10084			
10	10108–10536			
RP Thy 1	10988–11656	M27713—Thymidylate synthase complementing protein (Thy 1) of *Dictyostelium discoideum*; expect 3e-22 Possible hits to gp48 of mycobacteriophage D29/L5 and to gp16 of bacteriophage phi-C31	33	48
12	11623–11967			
13	12242–12610			
RP Primase	13380–14948	X17255—DNA primase/helicase of coliphage T3; expect 3e-74 Other significant hits included coliphage T7, bacteriophage SPP1, bacteriophage P1, *Mycoplasma arthritidis* bacteriophage MAV1, bacteriophage P22, and bacteriophage HK022	33	50
RP DNA Pol	15510–17244	X17255—DNA polymerase of coliphage T3; expect 2e-51 Also a significant hit for coliphage T7	26	41
16	17269–17847			
17	18273–19136			
RP Endodeoxyribonuclease I	19126–19530	V01146—Endodeoxyribonuclease 1 of coliphage T7; expect 4e-29 Also a significant hit for coliphage T3	48	72
19	19773–20006	Possible hit to gp37 of mycobacteriophage L5		
RP Ribonucleotide Reductase	20182–22197	L20047—Ribonucleoside triphosphate reductase of *Lactobacillus leishmannii*; expect 7e-36 Significant hits to gp50 of mycobacteriophage D29 and L5	26	40
RP PhoH	22526–23680	D10391—Phosphate starvation–inducible protein (PhoH/PsiH) of *Escherichia coli*; expect 5e-38	38	63
/22	24566–24904			
/23	25076–25678			
/24	25678–26421			
/25	26609–28303	Possible hit to bacteriophage N15 gp16		
/26	28369–32823			
/26.1	32838–33272			
/26.2	33274–33735			
/27	33710–35857			
/28	36130–36879			
/29	36881–37453			
/29.1	37571–38512			
/30	39292–38525			

SIO1. After optimization, the PCR protocol could consistently detect fewer than 10 copies of the Roseophage SIO1 genome directly in seawater or MSM samples (Fig. 5). The target resolution of PCR was verified using two methods. The PCR protocol was originally optimized using a plasmid that contained the region of Roseophage SIO1's genome to which the primer had been designed. DNA concentrations were used to determine the number of copies of this plasmid per PCR reaction. Once the PCR protocol had been optimized, Roseophage SIO1 lysates of a known titers were an-

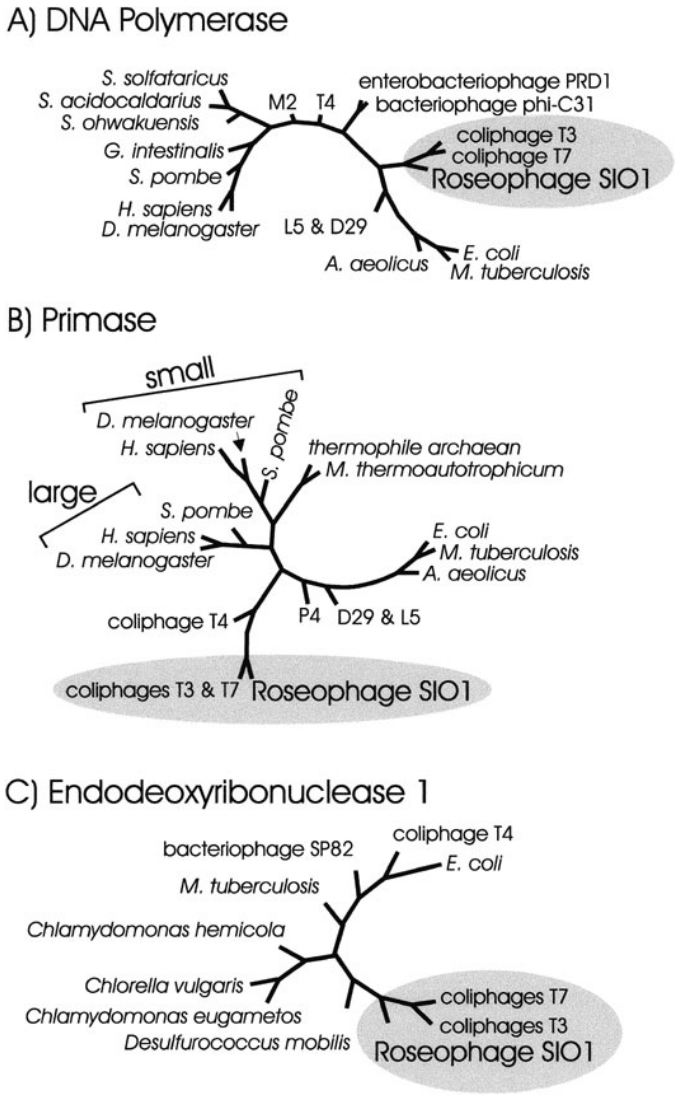

Fig. 4. Cladograms of (A) DNA polymerase, (B) primase, and (C) endodeoxyribonuclease 1. DNA polymerase A was used for the bacterial, archaeal, and eukaryotic specimens. Abbreviated phage are: Methanobacterium phage psi M2, coliphage T4, mycobacteriophage D29 & L5, and bacteriophage P4. Genus and species for abbreviated organism names are: *Escherichia coli, Mycobacterium tuberculosis, Aquifex aeolicus, Sulfolobus solfataricus, Sulfolobus acidocaldarius, Sulfurisphaera ohwakuensis, Methanobacterium thermoautotrophicum, Giardia intestinalis, Schizosaccharomyces pombe, Drosophila melanogaster,* and *Homo sapiens.*

alyzed in PCR reactions run parallel with the experiments. An example of the PCR controls is shown in Fig. 5, where 100, 10, and 0 refer to the number of phage per PCR reaction as determined by titering. Using this PCR protocol, it was shown that distinct Roseophage populations coexist at the same time and place and that Roseophage SIO1 was present in the ocean near Scripps Pier 8 yr after the original isolation. Interestingly, none of the Puget Sound isolates contained the Roseophage SIO1 sequence, suggesting geographical isolation of various Roseophages. The phage samples shown in Fig. 5 contained greater than 1,000 targets per reaction, as determined by titering. The Roseophage SIO1 isolated in 1997 (Roseophage SIO1b) was subsequently tested with two other Roseophage SIO1–specific primer sets. These primer sets gave identically sized products in both Roseophage SIO1 and SIO1b. Sequencing of the PCR products showed that 6 bp out of a total of 1,457 bp (0.4%) differed between Roseophage SIO1 and SIO1b.

Discussion

Genomic analysis allows the question of evolutionary relationships between marine and nonmarine phages to be addressed. Data presented in this study demonstrate that Roseophage SIO1 and coliphages T3 and T7 share a common ancestor for parts of the DNA replication machinery (i.e., primase, DNA pol, and endodeoxyribonuclease I). The mosaic model of phage evolution suggests that phage genomes consist of regions that are rearranged in different combinations to form new ''species'' (Goyal et al. 1987; Calendar 1988). Recently, Hendrix et al. (1999) reexamined the question of phage evolution using newly acquired genomic sequences. These authors convincingly show that dsDNA phage and prophage are mosaics that arose by horizontal transfer of genetic material from a global phage pool. The only restriction to the exchange of information is the frequency at which one phage can interact with another, and this is controlled in large part by host ranges. Phage overcome this restriction by small leaps in host range (i.e., a ''random walk through phylogenetic space,'' according to Hendrix et al. 1999). The finding that Roseophage SIO1 contains sequences highly similar to coliphages T3 and T7 supports this model, as does a recent finding that a number of marine cyanophages encode a region significantly similar to a capsid assembly protein (gp20) of coliphage T4 (Fuller et al. 1998). The trees shown in Fig. 4 indicate that many phage proteins (i.e., primase, DNA pol, and endodeoxyribonuclease I) are more closely related to other phage proteins than to host homologs. Since RP Thy1, RP Reductase, and RP PhoH are more closely related to nonphage proteins, these sequences may have been acquired via recombination with their hosts. However, this scenario is harder to envision for RP Thy1, which is most similar to a protein from the slime mold *Dictyostelium discoideum.* We favor the hypothesis that these proteins may be common components of other as yet uncharacterized phage. As more phage genomes are sequenced, the evolutionary relationships among them will become clearer. Assuming that phage evolve according to the mosaic model, the question arises, how long does a specific arrangement of ''phage parts'' persist in the environment? For Roseophage SIO1, it appears that this genomic arrangement has survived in the coastal water off La Jolla, California, for at least 8 yr (Fig. 5).

The basic life histories of marine phages can be elucidated by comparison of complete genomes to those of other extensively studied phages (e.g., lambda, T4, and T7). For example, the DNA replication machinery of Roseophage SIO1 shows a clear homology with that of coliphages T3 and T7, suggesting that the process of DNA replication may be similar among these phages. Other interesting clues about Ro-

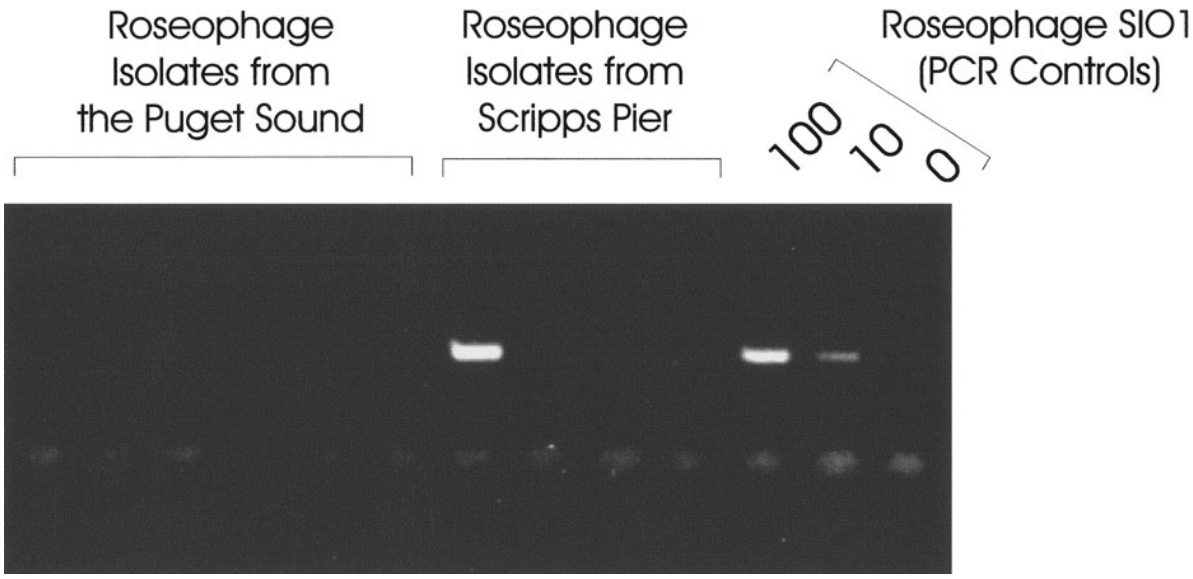

Fig. 5. Detection of Roseophage SIO1 in the environment. Phage that lytically infect *Roseobacter SIO67* were isolated from the Puget Sound and Scripps Pier water samples by plaque purification. On the final round of plaque purification, the agar plug was placed in MSM buffer and titered. The titers of all experimental samples were greater than 1,000 phage L^{-1}. PCR was then performed to detect phage containing the same sequence as Roseophage SIO1. The PCR controls were samples of known titers and contained 100, 10, and 0 Roseophage SIO1.

seophage SIO1's life history come from the absence of certain expected protein regions. For example, there are no regions of the Roseophage SIO1 genome that encode recognizable phage structural proteins. This is somewhat surprising because the morphology of Roseophage SIO1 (i.e., the capsid, tail, and nucleocapsid shown in Fig. 2) are very similar to that of other Podoviruses. Similarly, it is conspicuous that the Roseophage SIO1 genome lacks a recognizable RNA polymerase. Upon infecting *Escherichia coli*, coliphages T3 and T7 depend on the host's RNA polymerase for the transcription of Class I (Early) genes (Calendar 1988). One of the Class I gene products is a phage-encoded RNA polymerase responsible for transcribing the rest of the viral genome. The absence of a RNA polymerase in Roseophage SIO1 suggests two possibilities: the sequence encoding the RNA polymerase is so divergent from previously sequenced RNA polymerases that it is unrecognizable, or Roseophage SIO1 uses the host RNA polymerase to transcribe all of its genes. Genomic analysis favors the second possibility, since RNA polymerases contain a number of highly conserved regions that should be recognized by the algorithms used in these studies (Klenk et al. 1994). There is also evidence suggesting that transcriptional control of Roseophage SIO1's genes may be tied into the host's transcriptional networks. Upstream of RP PhoH is a transcriptional element known as the Pho-Box (data not shown). In *Escherichia coli*, phosphate starvation induces Pho B to bind to the Pho-Box, thereby activating the transcription of numerous genes (Neidardt 1987; Wanner 1996). The conservation of the Pho-Box in Roseophage SIO1 implies that expression of the phage's genes may be connected to phosphate metabolism.

The Roseophage SIO1 genome has at least four proteins involved in phosphate metabolism: RP PhoH, RP ribonucleotide reductase, RP Thy1, and RP endodeoxyribonuclease I. Two of these, the reductase and endodeoxyribonuclease I, have been found in a number of phage genomes (Calendar 1988; Hendrix et al. 1999). However, this is the first report of a phage carrying PhoH- or Thy1-like proteins. Although the incorporation of nucleotides from degraded host DNA has been well characterized in various phage–host systems (Calendar 1988), such recycling may be particularly important for marine phages whose hosts are much smaller and are growing in nutrient-poor environments. Experiments with Roseophage SIO1 and two other marine phages have shown that these phage use host nucleotides during phage replication (Wikner et al. 1993; in this reference, Roseophage SIO1 is referred to as MΦ14/67).

Many of the predictions about Roseophage SIO1's life history, derived from the genomic data, should be experimentally confirmed. However, given that there are probably millions of marine phage, detailed experimental analysis will not be practical for even a small subset of the marine phage. Ideally, we need to expand our knowledge about the evolutionary relationships and unique characteristics of marine phage before analyzing details of their life histories. Genomic analysis presents an alternative to time-consuming and expensive laboratory studies, especially since sequencing continues to become cheaper and faster. Additionally, there is a very large literature describing many different phages that can be used as models (much like *E. coli* is used as a bacterial model of cellular physiology and biochemistry). Genomic analysis will allow the identification of important differences between the model phages and marine phages and will aid in the design of laboratory experiments to address these differences.

References

ALTSCHUL, S. F., T. L. MADDEN, A. A. SCHÄFFER, J. ZHANG, Z. ZHANG, W. MILLER, AND D. J. LIPMAN. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. **25:** 3389–3402.

AZAM, F. 1998. Microbial control of oceanic carbon flux: The plot thickens. Science **280:** 694–696.

BERGH, O., K. Y. BØRSHEIM, G. BRATBAK, AND M. HELDAL. 1989. High abundance of viruses found in aquatic environments. Nature **340:** 467–469.

BOHANNAN, B. J. M., AND R. E. LENSKI. 1997. Effect of resource enrichment on a chemostat community of bacteria and bacteriophage. Ecology **78:** 2303–2315.

BØRSHEIM, K. Y. 1993. Native marine bacteriophage. FEMS Microbio. Ecol. **102:** 141–159.

CALENDAR, R. 1988. The Bacteriophages. Plenum Press.

CORPET, F. 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. **16:** 10881–10890.

FENCHEL, T. 1986. The ecology of heterotrophic microflagellates. Adv. Microbial Ecol. **9:** 57–97.

FUHRMAN, J., AND R. T. NOBLE. 1995. Viruses and protists cause similar bacterial mortality in coast seawater. Limnol. Oceanogr. **40:** 1236–1242.

FUHRMAN, J. A. 1999. Marine viruses and their biogeochemical and ecological effects. Nature **399:** 541–548.

———, AND L. CAMPBELL. 1998. Microbial microdiversity. Nature **393:** 410–411.

———, AND C. A. SUTTLE. 1993. Viruses in marine planktonic systems. Oceanography **6:** 51–63.

FULLER, N. J., W. H. WILSON, I. R. JOINT, AND N. H. MANN. 1998. Occurrence of a sequence in marine cyanophages similar to that of T4 g20 and its application to PCR-based detection and quantification techniques. Appl. Environ. Microbiol. **64:** 2051–2060.

GONZÁLEZ, J. M., AND M. A. MORAN. 1997. Numerical dominance of a group of marine bacteria in the alpha-subclass of the class Proteobacteria in coastal seawater. Appl. Environ. Microbiol. **63:** 4237–4242.

GOYAL, S. M., C. P. GERBA, AND G. BITTON. 1987. Phage ecology. Wiley.

HAWKSWORTH, D. L., AND M. T. KALIN-ARROYO. 1995. Magnitude and distribution of biodiversity, p. 107–192. *In* V. H. Heywood and R. T. Watson [eds.], Global biodiversity assessment, Cambridge Univ. Press.

HENDRIX, R. W., M. C. M. SMITH, R. N. BURNS, M. E. FORD, AND G. F. HATFULL. 1999. Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. Proc. Natl. Acad. Sci. USA **96:** 2192–2197.

HIDAKA, T., AND K. ICHIDA. 1976. Properties of a marine RNA-containing bacteriophage. Mem. Fac. Fish. Kagoshima Univ. **25:** 77–89.

KELLOGG, C. A., J. B. ROSE, S. C. JIANG, J. M. TURMOND, AND J. H. PAUL. 1995. Genetic diversity of related vibriophages isolated from marine environments around Florida and Hawaii, USA. Mar. Ecol. Prog. Ser. **120:** 89–98.

KLENK, H-P., P. PALM, AND W. ZILLIG. 1994. DNA-dependent RNA polymerases as phylogenetic marker molecules. Sys. Appl. Microbiol. **16:** 638–647.

LENSKI, R. E. 1988. Dynamics of interactions between bacteria and virulent bacteriophage. Adv. Microb. Ecol. **10:** 1–44.

LEVIN, B. R., F. M. STEWART, AND L. CHAO. 1977. Resource-limited growth, competition, and predation: A model and experimental studies with bacteria and bacteriophage. Am. Nat. **111:** 3–24.

MIDDELBOE, M., N. O. G. JORGENSEN, AND N. KROER. 1996. Effects of viruses on nutrient turnover and growth efficiency of noninfected marine bacterioplankton. Appl. Environ. Microbiol. **62:** 1991–1997.

MILLER, R. V. 1998. The importance of the dynamics of bacteriophage-host interactions to bacterial abundance and genetic diversity in aquatic environments. EPA Environ. Res. Brief.

MOEBUS, K. 1992. Further investigation on the concentration of marine bacteriophages in the water around Helgoland with reference to the phage-host systems encountered. Helgol. Meeresunters. **46:** 275–292.

———, AND H. NATTKEMPER. 1991. Bacteriophage sensitivity patterns among bacteria isolated from marine waters. Helgol. Meeresunters. **34:** 375–385.

MURPHY, F. A., D. H. L. BISHOP, S. A. GHABRIAL, A. W. JARVIS, G. P. MARTELLI, M. A. MAYO, AND M. D. SUMMERS. 1995. Virus taxonomy: Sixth report of the international committee on taxonomy of viruses. Springer-Verlag.

NEIDHARDT, F. C. 1987. *Escherichia coli* and *Salmonella typhimurium* cellular and molecular biology. Am. Soc. Microbiol.

PINHASSI, J., U. L. ZWEIFEL, AND A. HAGSTROM. 1997. Dominant marine bacterioplankton species found among colony-forming bacteria. Appl. Environ. Microbiol. **63:** 3359–3366.

PROCTOR, L. M., AND J. A. FUHRMAN. 1990. Viral mortality of marine bacteria and cyanobacteria. Nature **343:** 60–62.

SAMBROOK, J., E. F. FRITSCH, AND T. MANIATIS. 1989. Molecular cloning, a laboratory manual. Cold Spring Harbor Press.

SCHMIDT, T. M., E. F. DELONG, AND N. R. PACE. 1991. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. J. Bacteriol. **173:** 4371–4378.

SHERR, B. F., AND E. B. SHERR. 1984. Role of heterotrophic protozoa in carbon and energy flow in aquatic ecosystems, p. 412–423. *In* M. J. Klugg and C. A. Reddy [eds], Current perspectives in microbial ecology. ASM Press.

SMITH, D. C., M. SIMON, A. L. ALLDREDGE, AND F. AZAM. 1992. Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution. Nature 359: 139–142.

STEWARD, G. F. , AND F. AZAM. In press. Analysis of viral assemblages. *In* C. R. Bell, M. Brylinsky, and P. Johnson-Green [eds.] Microbial biosystems: New frontiers. Atlantic Canada Society for Microbial Ecology, Halifax.

———, D. C. SMITH, AND F. AZAM. 1996. Abundance and production of bacteria and viruses in the Bering and Chukchi Seas. Mar. Ecol. Prog. Ser. **131:** 287–300.

———, J. WIKNER, W. P. COCHLAN, D. C. SMITH, AND F. AZAM. 1992. Estimation of virus production in the sea: II. Field results. Mar. Microb. Food Webs **6:** 79–90.

SUTTLE, C. A. 1993. Enumeration and isolation of viruses, p. 121–134. *In* P. F. Kemp, B. F. Sherr, E. B. Sherr, and J. J. Cole [eds], Handbook of methods in aquatic microbial ecology. Lewis Publishers.

———. 1994. The significance of viruses to mortality in aquatic microbial communities. Microb. Ecol. **28:** 237–243.

———, AND A. M. CHAN. 1993. Marine cyanophage infecting oceanic and coastal strains of *Synechococcus*: Abundance, morphology, cross-infectivity, and growth characteristics. Mar. Ecol. Prog. Ser. **92:** 99–109.

———, AND ———. 1994. Dynamics and distribution of cyanophages and their effect on marine *Synechoccus* sp. Appl. Environ. Microbiol. **60:** 3167–3174.

WANNER, B. L. 1996. Phosphorus assimilation and control of the phosphate regulon, p. 1357–1381. *In* F. C. Neidhardt [ed.] *E. coli* and *Salmonella*: Cellular and molecular biology. ASM Press.

WATERBURY, J. B., AND F. W. VALOIS. 1993. Resistance to co-occurring phage enables marine *Synechococcus* communities to

coexist with cyanophages abundant in seawater. Appl. Environ. Microbiol. **59:** 3393–3399.

WIKNER, J., J. J. VALLINO, G. F. STEWARD, D. C. SMITH, AND F. AZAM. 1993. Nucleic acids from the host bacterium as a major source of nucleotides for three marine bacteriophages. FEMS Microbiol. Ecol. **12:** 237–248.

WOMMACK, K. E., J. RAVEL, R. T. HILL, J. CHUN, AND R. R. COLWELL. 1999. Population dynamics of Chesapeake Bay virioplankton: Total-community analysis by pulsed-field gel electrophoresis. Appl. Environ. Microbiol. **65:** 231–240.

Received: 9 July 1999
Amended: 5 October 1999
Accepted: 14 October 1999